

Differential Effect of Reward and Punishment on Procedural Learning

Tobias Wächter,^{1,2,3} Ovidiu V. Lungu,^{1,2} Tao Liu,^{1,2} Daniel T. Willingham,⁴ and James Ashe^{1,2,3}

¹Brain Sciences Center, Veterans Affairs Medical Center, Minneapolis, Minnesota 55417, Departments of ²Neuroscience and ³Neurology, University of Minnesota, Minneapolis, Minnesota 55455, and ⁴Department of Psychology, University of Virginia, Charlottesville, Virginia 22904

Reward and punishment are potent modulators of associative learning in instrumental and classical conditioning. However, the effect of reward and punishment on procedural learning is not known. The striatum is known to be an important locus of reward-related neural signals and part of the neural substrate of procedural learning. Here, using an implicit motor learning task, we show that reward leads to enhancement of learning in human subjects, whereas punishment is associated only with improvement in motor performance. Furthermore, these behavioral effects have distinct neural substrates with the learning effect of reward being mediated through the dorsal striatum and the performance effect of punishment through the insula. Our results suggest that reward and punishment engage separate motivational systems with distinctive behavioral effects and neural substrates.

Key words: procedural learning; basal ganglia; reward; punishment; sequence; motor; motivation

Introduction

Reward and punishment are potent modulators of human and animal behavior (Thorndike, 1911; Pavlov, 1927; Skinner, 1938; Sutton and Barto, 1998). However, despite the great increase in knowledge in the past two decades of the neural basis of the reward effect (Schultz, 2002), and that of punishment to a lesser extent, we lack clear data about how reward and punishment influence the learning of specific behaviors, apart from those in classical and instrumental conditioning, and how this might be mediated at a neural level (Delgado, 2007). To address this issue, we focused on procedural learning, a distinct learning behavior that is the foundation of many of our motor skills and perhaps other functions such as cognitive, category, and perceptual learning (Squire, 2004). The fact that procedural learning is thought to be mostly dependent on the basal ganglia (Willingham et al., 2002), which also mediates the effect of reward and punishment (Schultz, 2002), makes it an ideal behavior to study.

To test the influence of reward and punishment on procedural learning, we used a modified version of the serial reaction time (SRT) task (Nissen and Bullemer, 1987), a simple and robust experimental probe (Robertson, 2007), during which continuous modulation of motor output was required and reinforcement and nonreinforcement learning were dissociated. We opted to use monetary reward (and punishment) because it is a strong modulator of human behavior and has clear effects on brain activity (Breiter et al., 2001; Delgado et al., 2003). As in the original

SRT task, subjects pressed one of four buttons with the right hand when instructed by a visual display. Trials were presented in blocks in which the lights were illuminated either randomly or, unbeknownst to the subject, on the basis of a 12-element repeating sequence. All subjects first performed several blocks of random trials to minimize the effects of learning the general visuo-motor behavior on later blocks and to establish an individual criterion response time (cRT) on which subsequent reward and punishment would be based. Subjects were then randomly assigned to reward, punishment, and control groups. Because we used reaction time (RT) as an index of learning, subjects were, therefore, rewarded for faster responses, which were easiest to generate through learning the repeating sequence. To control for a potential distractor effect of the incentives on the learning process, several additional blocks without reward or punishment were presented after the trials with incentives. The blocks were presented in the order outlined in Figure 1A,B. In the current report, we describe the impact of reward and punishment behavioral results using a large cohort of subjects and the neural correlates of this behavior in a smaller number of different subjects on whom functional imaging was also performed.

Materials and Methods

Subjects

In the behavioral study, we tested 91 healthy human subjects (69 female) (mean, 21.7 ± 3.5 years); in the functional imaging component, we studied 41 right-handed subjects (22 female; 21 years ± 2.58). All subjects were of similar socioeconomic backgrounds recruited from the University community and gave informed consent to participate in the study, which was approved by the local Institutional Review Board.

Behavioral task

The task (see Fig. 1) was a modification of the original SRT task of Nissen and Bullemer (1987). Subjects were presented with four visual stimuli arranged horizontally. Each stimulus was associated with one of the four fingers of the right hand; illumination of the visual stimulus indicated the

Received Aug. 28, 2008; revised Nov. 19, 2008; accepted Nov. 29, 2008.

This work was supported by National Institutes of Health (NIH) Grant NS40106 (J.A. and D.T.W.), NIH Center Grant P41RR008079, and the Department of Veterans Affairs. We thank Bruce Overmier, Mary Jo Nissen, John O'Doherty, and Nima Khoshnoodi for comments on a previous version of this manuscript.

Correspondence should be addressed to James Ashe, Brain Sciences Center (11B), Veterans Affairs Medical Center, One Veterans Drive, Minneapolis, MN 55417. E-mail: ashe@umn.edu.

DOI:10.1523/JNEUROSCI.4132-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/290436-08\$15.00/0

finger to be moved. Subjects used a key-press device to respond as quickly as possible when instructed by the visual display. Stimuli were presented in blocks of 96 trials in either a pseudorandom or repeating 8- × 12-element sequence with an 800 ms interstimulus interval. To minimize awareness, the sequence was constrained (Willingham et al., 2000). Contrary to the original SRT task, random and sequence blocks were presented both with (R-blocks and S-blocks) or without (r-blocks and s-blocks) monetary incentives, respectively. For the purposes of the current experiment, we define reward and punishment in terms of positive and negative monetary incentives.

Of the 91 original subjects, 25 subjects were excluded because they developed explicit knowledge of the sequence (see below), and a further two subjects were eliminated because they did not perform the test properly, leaving a cohort of 64 subjects. At the beginning of the experiment, subjects performed four practice r-blocks; we then calculated a cRT for each subject based on their median RT in the last of the four blocks. Subjects were then randomized into reward ($n = 21$), punishment ($n = 24$), and control ($n = 19$) groups. It is important to note, that because learning occurs implicitly in the SRT task, reward or punishment could not be applied directly in this paradigm. However, learning in the SRT task leads to faster reaction time, and this is usually used as a behavioral index of learning in this task. We therefore rewarded/punished subjects according to the change (increase/decrease) of their individual reaction time. Those in the reward group were informed that they would be rewarded (+ 4 cents) for each trial for which their RT was less than cRT; the penalty group were penalized (− 4 cents) if their RT was greater than their cRT. Rewarded subjects started with \$0, whereas punished group were given \$38. Because of normalization of incentives to base performance, both rewarded and punished subjects ended up with an average of \$21–22, whereas the control subjects received a fixed payment of \$23. The incentive schedule also controlled for motivation in the two groups. Subjects received ongoing feedback of their performance both through an incrementing (or decrementing) counter displaying their current monetary position displayed on the screen and also by the color of the visual stimuli (green and red stimuli indicated that the RT on the preceding trial was less or greater than the cRT, respectively). Because learning of the sequence enabled subjects to reduce their RT significantly compared with the cRT, those in the reward group were rewarded for learning, whereas the penalty group were punished if they did not learn. Control subjects were presented with an equal number of red and green stimuli and told that the color was of no significance. To control for the potential distractor effect of the counter in the incentive groups, the controls saw an identical tally that kept a count of either red or green lights. After the initial practice session (four consecutive r-blocks), the full experiment comprised 15 blocks in the following order: r-R-R-R-R-S-S-S-S-R-r-r-s-r (r , random blocks; S = sequence blocks; upper case, blocks with incentives in rewarded and punished groups; lower case, blocks without incentives). The color of the feedback visual stimuli was counter-balanced across subjects (supplemental Fig. S1, available at www.jneurosci.org as supplemental material). Blocks were separated by 30 s breaks. Of the 64 subjects that are the basis of the behavioral component of the current report, the first 27 subjects performed only the first 11 blocks; the remaining four blocks were added later to measure the knowledge transfer in the absence of incentives.

A different group of 41 (reward, $n = 11$; punished, $n = 13$; control, $n = 17$) subjects was used for the imaging experiment of whom 32 (reward, $n = 11$; punished, $n = 13$; control, $n = 17$) learned the sequence implicitly. The behavioral paradigm was identical to that of the behavioral experiment, with the exception of the monetary incentives which were adjusted as follows: the reward group started with \$0 and was rewarded with 10 cents for every correct response; in the punished group, 20 cents were deducted for each incorrect response from a base of \$95; the average final payment for the rewarded and punished groups was \$55, which was also the fixed payment given to the control group.

Immediately after performing both the behavioral and the imaging experiments, all subjects were tested for explicit knowledge of the sequence. The participants were told that the stimuli might have appeared in a repeating sequence and were asked to reproduce the sequence in a free recall task. They were allowed to use the key-press device if they

thought it might be helpful. Subjects who could not recall anything were encouraged to guess. If a participant could reproduce a correct string of >4 consecutive items of the sequence, learning was considered to be explicit and the subject excluded from the analysis. We used a relatively conservative measure of explicit knowledge (Willingham and Goedert-Eschmann, 1999) to ensure that the effects described were primarily attributable to procedural knowledge. To limit the number of subjects excluded from the functional magnetic resonance imaging (fMRI) analysis because of explicit learning, we used a slightly less conservative measure in this component of the experiment, and the learning was considered to be explicit if a subject could reproduce a correct string of >5 consecutive items of the sequence.

fMRI acquisition and preprocessing

Imaging parameters. Image acquisition was done with a 3 Tesla whole-body MR system (MAGNETOM Trio; Siemens Medical Systems) using a Siemens head coil. High-resolution (1 mm iso-voxel) T1-weighted-3d-FLASH sagittal sequences [160 slices per slab; field of view (FOV), 256 × 256 mm, repetition time (TR), 20 ms; echo time, 4.7 ms; flip angle, 22°; slice thickness, 1 mm] was first acquired to enable localization of functional images. Thereafter, whole-brain fMRI was performed using an echoplanar imaging sequence measuring blood oxygenation level-dependent (BOLD) signal. A total of 30 functional slices per volume were acquired for all subjects in all runs. These slices had a thickness of 3 mm, and they were acquired in the transverse plane (matrix size, 64 × 64) FOV 192 × 192 mm, with a 33% gap. A complete scan of the whole brain was acquired in 2560 ms (TR), flip angle 80°, TE = 30 ms, and a total number of 611 volumes were acquired for the whole experiment.

Preprocessing of fMRI data. Brain Voyager (Brain Innovation B.V.) software was used for fMRI data preprocessing and analysis. The functional bidimensional images of every subject were preprocessed to correct for motion artifacts (movements <3 mm in any plane), for differences in slice scan time acquisition and for temporal linear trends. Then, these functional images were used to reconstruct the three-dimensional (3D) functional volume for every subject and every run. Next, we performed spatial smoothing of the functional data using a Gaussian full-width at half-maximum kernel of 7 mm. The 3D functional volume was subsequently aligned with the corresponding 3D anatomical volume, and both of them were normalized to standard Talairach space.

Statistical analyses

Behavioral data. Statistical analyses were done using SPSS 11.0. Response time gains in individual subjects were calculated by deducting the RT of each trial from the cRT for that subject. Because RT gains were negative values and not normally distributed, we added 1000 to each value and did a logarithmic transformation before any statistical calculations. However, when differences in RT or RT gains are stated or when displayed in graphs, the original, nontransformed scale was used. Generalized linear models (GLM) were used for comparison of different conditions among groups. The threshold for significance was 0.05 adjusted for multiple comparisons, if necessary. To determine the extent of learning in each of the three behavioral groups (reward, punishment, control), we used a custom orthogonal contrast between the sequence block (14) and surrounding random blocks (13 and 15) for which we used the following contrast weights: $\frac{1}{2}$, -1 , $\frac{1}{2}$ for blocks 13, 14, and 15, respectively.

fMRI data. GLM models were used for voxelwise data analysis. A block design protocol was used to demonstrate the main effects. To determine the “effect of procedural learning” in all subjects we regrouped the 15 blocks into four regressors [random and sequence, with (R,S) and without (r,s) incentive] and the baseline in rewarded and punished subject groups. To address the neural basis of the “performance effect” we did an additional block-based analysis using a GLM that included each block (total 15) as an independent regressor and contrasted the first nonincentive block (r ; block 1) with the first incentive block (R ; block 2). An event-related protocol was used to assess the influence of incentives by testing the difference between trial types (rewarded, punished, control) and differences within types relating to the presence or absence of incentive. The event-related analysis was performed for blocks 2–11, and each behavioral event (800 ms) was classified according to our experimental

manipulation (reward, punishment, or control) and the performance of the subject (green or red lights). These predictors were entered as fixed factors in a mixed GLM model, with subjects as random factors (to control for possible differences among subjects).

The statistical parameters of these models were computed voxelwise for the entire brain, and activation maps were computed for various contrasts between the predictors. The criteria used for activation maps we generated were as follows: cluster size of 10 adjacent voxels (to demonstrate the performance effect, the cluster size was raised to 100 adjacent voxels) and a statistical threshold for the cluster of $p < 0.001$. The output of these models was selectively used for region of interest analysis both of the learning and the performance effects using SPSS.

We performed two additional voxelwise analyses. The first analysis (supplemental Table S3, available at www.jneurosci.org as supplemental material) was to determine the brain areas that related to the enhanced learning in the reward group (punished group as control). We first applied the main effects of the block design model as a mask (uncorrected, $p < 0.05$; voxels $> 108 \text{ mm}^3$) within which we tested the interaction between learning and group ($t = 3.174$; $p < 0.005$; voxels $> 108 \text{ mm}^3$); group was defined as reward/punishment and learning as sequence (S blocks)/baseline. The second analysis (supplemental Table S5, available at www.jneurosci.org as supplemental material) was to detect the functional activation associated with the performance effect in the punishment group. A GLM was built that included each block (total 15) as an independent regressor. The normalized cRT (mean, 0; SD, 1; only for first two blocks) was included in this model to determine the neural basis of the performance effect.

Results

As mentioned above, we first present the data from a behavioral study on a large cohort of subjects followed by the data from functional imaging on a different but smaller group performing the same task.

Behavioral effects of reward and punishment

In the behavioral study, we tested 91 subjects, of whom 25 developed explicit knowledge of the sequence and two did not perform the experiment properly and were excluded from further analyses. The remaining 64 subjects were divided among the reward ($n = 21$), punishment ($n = 24$), and control ($n = 19$) groups. Learning in the SRT task is generally assessed by RT gain, or in some cases by the error rate in sequence blocks compared with random blocks (Fig. 1). There were three main findings in the behavioral study. First, there was a strong non-learning-related performance effect in the punishment group. This effect was manifested by a sharp and significant (pairwise comparisons using Bonferroni test, $p < 0.001$) 19.6 ms drop in RT in the punishment group during random blocks immediately after the implementation of penalties (Fig. 1A, blocks 1–5). Because there was no sequence to be learned, the RT decrease must have been caused by a pure performance effect provoked by punishment and not related to learning. A comparable effect was not seen in the reward group. The second finding was a differential learning effect in the three groups (group \times block interaction; $F_{(8,44128)} = 3.39$; $p = 0.001$) during the learning phase (Fig. 1A, blocks 6–9). At the end of the sequence blocks, the reward and control groups showed a significantly greater (pairwise comparisons using Bonferroni test, $p < 0.001$) decrease in RT than the punishment group. However, the relatively modest RT gain in the punished group during sequence blocks may have been attributable to a floor effect as a result of the already low RT from improvement in performance in the preceding random blocks. To test for possible floor effects and also to eliminate potential distractor effects that might have unpredictable consequences on the

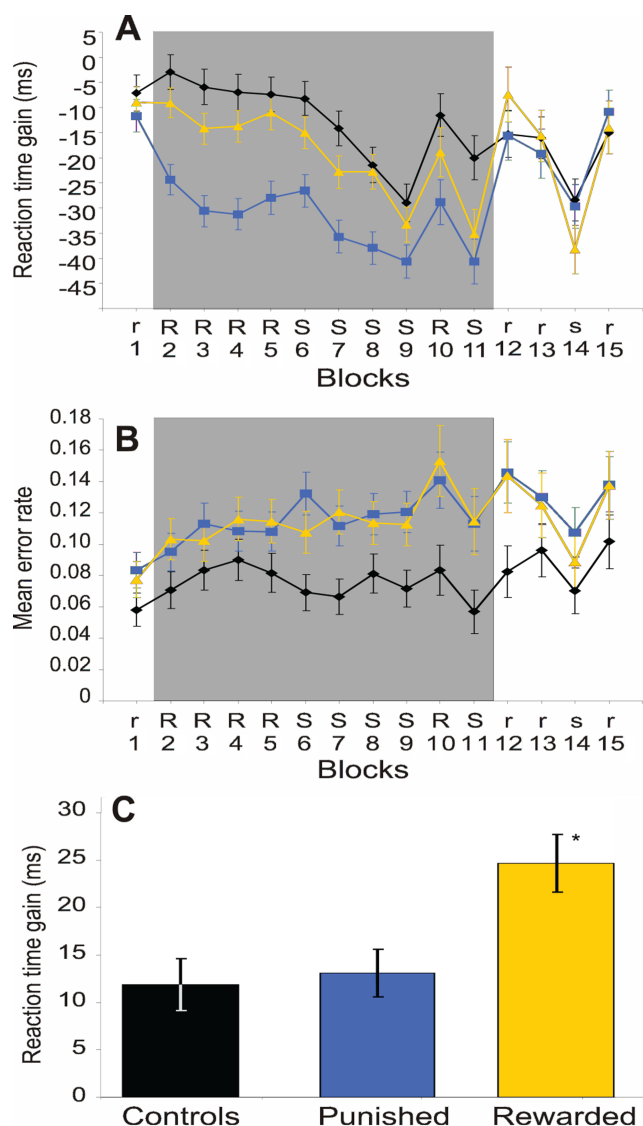


Figure 1. Behavioral results. **A**, **B**, Time course of response time gains (**A**) and error rates (**B**) in subject groups during task performance. The mean (\pm SEM) response time gain for each block, relative to individual subjects' cRT, in the reward (yellow; $n = 21$), punishment (blue; $n = 24$), and control (black; $n = 19$) groups. The gray hatching indicates the blocks during which incentives were used. Random and sequence blocks were presented with (R, S) and without (r and s) monetary incentives, respectively. **C**, The absolute gain in response time attributable to learning in subject groups. The mean (\pm SEM) of the absolute response gain in the transfer portion of the task (blocks 12–15) was derived by comparing RT in the sequence block (block 14) to the mean of the adjacent blocks (blocks 12, 13, and 16) for each group. The difference between the reward and the other two groups was significant ($p = 0.003$). Error bars indicate SEM.

expression of learning (Frensch et al., 1999), we added a transfer phase, without reward or punishment in 37 of the 64 subjects performing the task (Fig. 1A, blocks 12–15). Using these transfer blocks, sequence learning was estimated by a custom orthogonal contrast between the sequence block (14) and surrounding random blocks (13 and 15) for which we used the following contrast weights: $\frac{1}{2}$, -1 , $\frac{1}{2}$ for blocks 13, 14, and 15, respectively. The results of this contrast showed that learning occurred in all groups. However, the reward group learned more than either of the other groups (RT savings: reward, 25.9 ms; control, 13.0 ms; punishment, 14.5 ms), and this difference (Fig. 1C) was significant [$F_{(2,12232)} = 6.07$; $p = 0.002$;

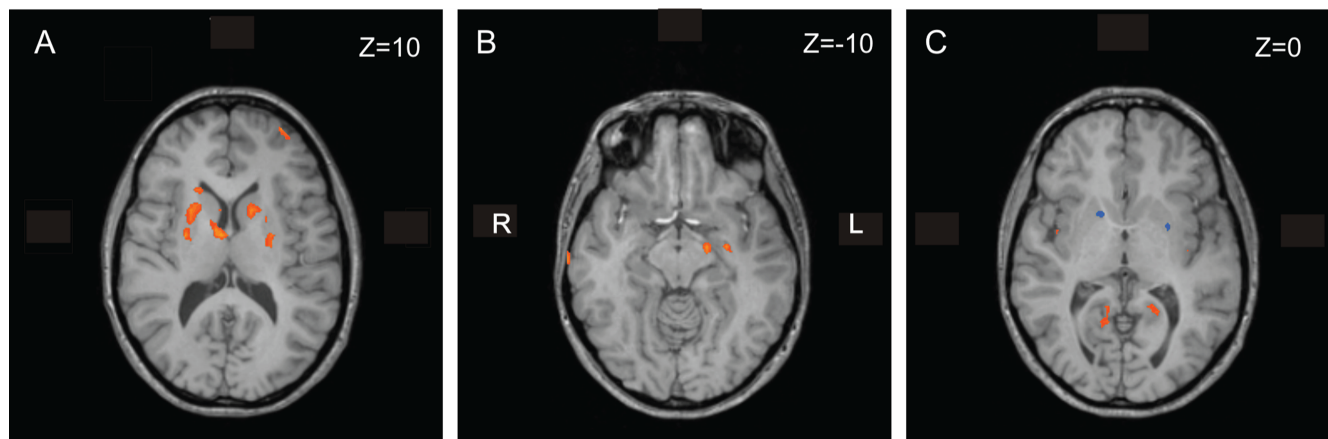


Figure 2. Effect of reward and punishment on brain activity. **A, B**, Event-related averaging in the rewarded group ($n = 11$); when rewarded trials (green stimuli) were compared with trials without reward (red stimuli), there was significant ($p < 0.001$) activation in the striatum on both sides as well as the nucleus accumbens on the left. **C**, Trials with punishment compared with trials without in the punished group ($n = 13$) result in a decreased BOLD signal in the striatum bilaterally (activation above baseline during neutral stimuli) and an increase in the insula bilaterally. There was no activation in these regions comparing red and green stimuli (both without incentives) in the controls ($n = 17$).

mean square error (MSE), 0.006], in which the 95% confidence interval of the rewarded group did not overlap with the confidence interval of either of the other two groups. Finally, the motivational effect of the incentives, based on a comparable speed–error trade-off profile (Fig. 1B), was similar in both groups but was different from that in the controls ($F_{(2,75795)} = 126.19$; $p < 0.001$; MSE, 0.092). The error rate, however, may also be used as an index of learning and, as expected, decreased during sequence blocks and showed the greatest decrease in the reward group in the transfer blocks (12–15), although this difference did not reach statistical significance.

The behavioral experiment, therefore, showed that only reward but not punishment enhances the implicit learning of sequences. This finding is at odds with previous studies of reward and punishment on learning, which, however, were generally confined to associative-learning tasks. This result provided us with an opportunity to test the neural substrate of the reward and punishment effects during procedural learning.

Neural substrates of learning and performance

We studied another group of naive subjects using fMRI during the performance of an almost identical behavioral task to the one described above and performed a detailed analysis on 41 subjects of whom 32 (reward, $n = 10$; punishment, $n = 11$; control, $n = 11$) did not acquire explicit knowledge of the sequence. Using data from the implicit learning cohort, we confirmed the previous behavioral results (supplemental Fig. S2, available at www.jneurosci.org as supplemental material). Specifically, we found a differential learning effect in the three groups (group \times block interaction; $F_{(6,11145)} = 2.42$; $p = 0.024$) during the learning phase. The greatest RT gain occurred in the rewarded, followed by the punished and control groups (pairwise comparisons using Bonferroni test, $p < 0.001$). The rewarded group had a gain of 27.12 ms, whereas the punished and control groups had significantly smaller, but similar gains (13.84 ms and 15.13 ms, respectively). These results suggest that the rewarded group learned more than either of the other groups when provided with incentives. There was a strong non-learning-related performance effect of incentive in the punished and rewarded groups when the random blocks with incentive were compared with those without

(interaction effect: $F_{(8,13966)} = 8.15$; $p < 0.001$). Similar to the purely behavioral study, the imaged subjects who were punished showed the greatest RT gain (29.16 ms) when compared with the other two groups (pairwise comparisons using Bonferroni test, $p < 0.001$). However, the rewarded group also showed significantly greater RT gains compared with the control group (17.14 ms vs 4.9 ms), albeit smaller than those of the punished group (pairwise comparisons using Bonferroni test, $p < 0.001$).

For the functional imaging analysis, we first examined the neural substrate of reward and punishment in all subjects over the whole experiment by comparing activation in the trials with and without incentives, separately for the reward and punishment groups (Fig. 2; supplemental Table S1, available at www.jneurosci.org as supplemental material) and focused primarily on subcortical activation. This contrast yielded an increase in activation in the dorsal striatum, in the nucleus accumbens, and amygdala, as well as in prefrontal cortex in the rewarded group (Fig. 2A,B). The punished group, as might be expected, showed quite a different pattern with increased activation during the punished trials in the insula bilaterally (Fig. 2C) and in portions of prefrontal cortex. Interestingly, the lack of punishment was associated with a relative increase in activation in the reward-related portions of the dorsal striatum. The comparison of activation associated with the randomly presented red and green stimuli in the control group showed no subcortical activation.

We next determined the areas responsible for procedural learning in all subjects (across groups) by comparing the activity during sequence blocks to the activity during random blocks. This analysis showed a significant increase in BOLD signal during the learning phase in the putamen bilaterally (Fig. 3A,B) as well as in the pontine gray matter and in the dentate nucleus on the ipsilateral side (supplemental Table S2, available at www.jneurosci.org as supplemental material) indicating that these areas are implicated in the procedural learning process. To evaluate the effect of incentives more specifically in these areas important for procedural learning, we plotted (Fig. 3C), for each implicit learning group separately, the changes in the BOLD signal in the putamen (left and right combined) during the following three trial types: (1) trials in which no incentive was possible (named “neutral” during r and s blocks), (2) incentive trials in R and S

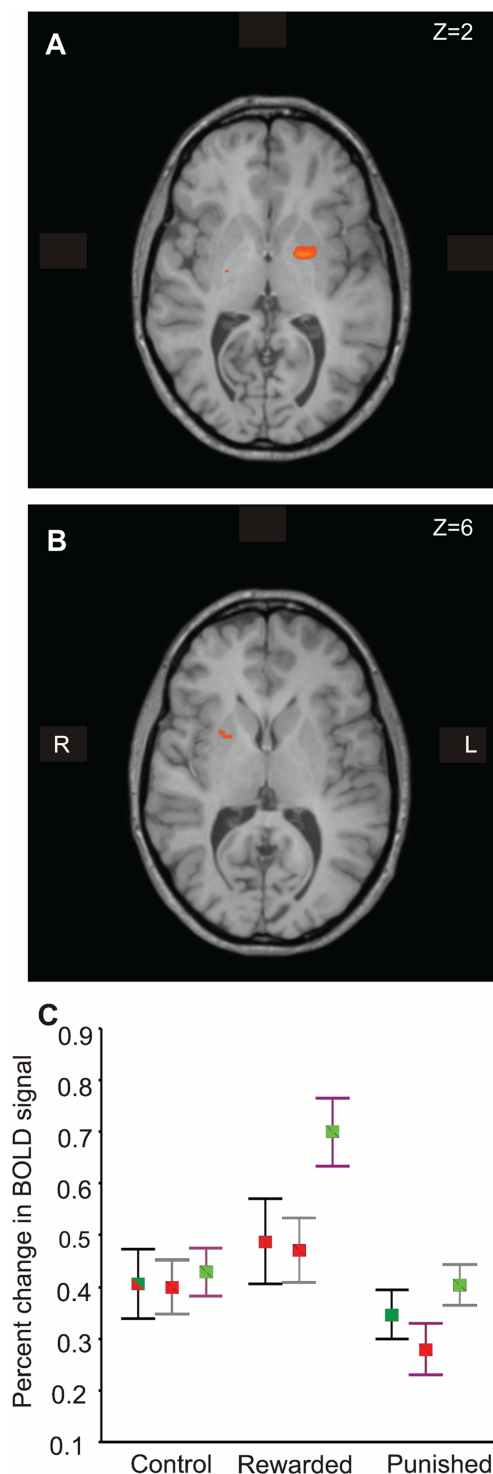


Figure 3. Activation in areas critical for procedural learning. **A, B,** A comparison of sequence blocks with random blocks over all groups in all subjects with implicit learning ($n = 41$) showed a significant ($p < 0.001$, uncorrected) increase in activation during motor learning in the corpus striatum (putamen and globus pallidus) bilaterally. **C,** More specific analysis of the striatal region in each group during implicit learning showed that in the rewarded group there was a significant ($p < 0.001$; Sidak tests adjusted for multiple comparisons) change in BOLD signal during trials in which subjects were rewarded (purple bar) compared with nonrewarded trials (gray bar) or trials in which no reward could be obtained (black bar). In contrast, the BOLD signal in the punished group during punished trials (purple bar) and nonpunished trials (gray bar) did not significantly diverge from that during trials in blocks where no punishment could be obtained (black bar). The difference between trials with punishment and trials without punishment is, however, significant ($p < 0.005$; Sidak tests adjusted for multiple comparisons). The trials in the control group show that there was no effect of stimulus color (square on each bar). The color of the visual stimulus for particular trial types is indicated at the midpoint of each bar. Error bars show the 95% confidence interval.

blocks (green trials for rewarded and red for punished group), and (3) nonincentive trials during R and S blocks (red trials for rewarded and green for punished group). We then analyzed the BOLD signal in these ROIs using a GLM model with group (control, rewarded, punished) and color of the stimulus (neutral, green, red) in the current trial as independent factors. Pairwise comparisons between various groups were assessed using Sidak tests, adjusting for the number of multiple comparisons. We found a significant main effect of group ($F_{(2,1016)} = 32.59$; $p < 0.001$; MSE, 0.093), color of the stimulus ($F_{(2,1016)} = 18.17$; $p < 0.001$; MSE, 0.093) and a significant interaction effect ($F_{(2,1016)} = 4.41$; $p < 0.001$; MSE, 0.093) between the independent factors. Sidak tests for pairwise comparisons showed that, overall, the BOLD signal change in the putamen was different in all three groups: it was the highest for the rewarded group, followed by controls and then by the punished group (all differences significant at $p < 0.05$, adjusted for multiple comparisons). In the rewarded group, there was a significant increase in BOLD signal (Sidak test significant at $p < 0.001$) for trials that were rewarded compared with nonrewarded trials or those in which no reward (incentive) was possible. In the punished group, the BOLD signal change during incentive trials (i.e., punishment; red stimuli) was not different from neutral trials (Sidak test, $p = 0.386$), but it was significantly less (Sidak test, $p < 0.005$) than that during nonincentive trials (green trials). It can be seen that the color of the stimulus had no effect on striatal activation in the control group. Furthermore, there was no significant difference in the BOLD signal during nonincentive trials between control and rewarded groups (red trials) or between control and punished groups (green trials), indicating a stable baseline activation when executing the task without incentives. We further confirmed the findings of this ROI approach using a voxel-based analysis of the interaction between learning and group (supplemental Table S3, available at www.jneurosci.org as supplemental material). The voxelwise analysis also showed an interaction effect in the middle frontal gyrus in an area consistent with the location of dorsal premotor cortex, which suggests that the effect we documented in basal ganglia is transmitted to the cortex thus facilitating changes in motor output.

Although in the behavioral experiment punishment showed no effect on the process of learning, it led to an immediate performance effect expressed by a decrease in reaction time immediately after the introduction of punishment (Fig. 1A) during the random blocks. To examine the neural substrate of this performance effect, we compared the BOLD response in the last block of random stimuli without incentives to the first block with incentives (supplemental Table S4, available at www.jneurosci.org as supplemental material). In the punished group, this contrast showed an increase in activation in the insula, and in the inferior frontal gyrus and hippocampus, in addition to other areas (Fig. 4). Among the areas activated in this contrast, only the change in activation in the insula was correlated (Pearson $r = -0.54$; $p < 0.005$; $n = 26$) with the performance change in punished subjects. Although the rewarded subjects also showed a performance effect, this did not correlate with changes in insula activation (Pearson $r = 0.001$; $p = 0.997$; $n = 22$). We confirmed the performance-related activation of the insula in a separate voxelwise analysis (supplemental Table S5, available at www.jneurosci.org as supplemental material).

Discussion

Our results point to fundamental differences between the effects of reward and punishment on behavior and to their quite

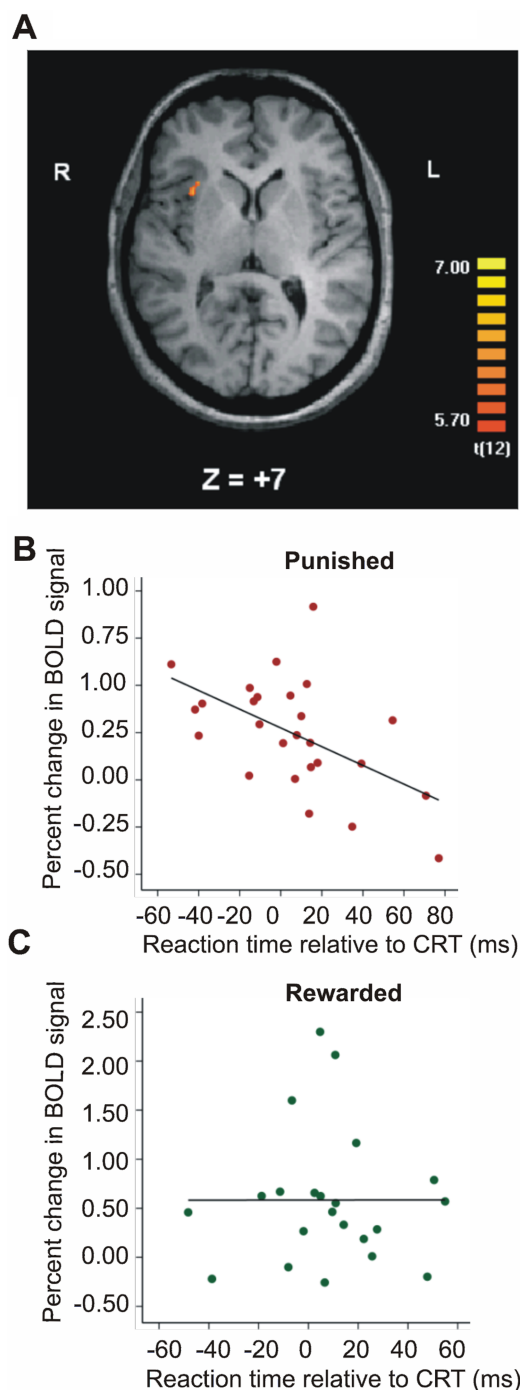


Figure 4. Neural substrate of change in performance. **A**, Activation in right insula, in punished subjects only, after the contrast between block 2 (R) random, with incentives (green and red stimuli combined), and block 1 random (r), without incentives. Statistical threshold was $t_{(12)} \geq 5.7$ with a minimum cluster size of 100 voxels ($1 \times 1 \times 1$ mm resolution). **B**, **C**, The correlation between changes in BOLD signal and reaction time relative to CRT for both blocks 1 and 2 combined for the rewarded (**B**) and the punished (**C**) groups. In each panel, each subject has a pair of points, one from block 1 and one from block 2.

distinct neural substrates. In addition, we extend the relevance of reward-based learning in the basal ganglia to the learning of habits (Packard and Knowlton, 2002) and skills and sequences (Houk et al., 1995; Berns and Sejnowski, 1998; Suri and Schultz, 1998).

The first issue to consider is whether reward and punishment have similar effects on behavior. We designed the oppo-

nent behaviors in our experiment as they might be constructed in every day life. Our objective was to encourage fast responses to the visual stimuli; therefore, we rewarded the desired behavior in one group and punished the undesired behavior in another. Both manipulations had a measurable effect on behavior, and deciding which is preferable depends on whether one is interested in short term changes in performance without enhancement of learning or longer term changes in learning itself. Only the reward group showed enhanced implicit learning of the motor sequence, although the punished group also learned they did not do any better than the control subjects. The lack of an effect of punishment on learning, although there were clear effects on other aspects of behavior (Fig. 1), was surprising and would not necessarily have been predicted on the basis of the reinforcement learning literature. B. F. Skinner regarded punishment as a “questionable technique,” speculated as to whether it actually worked, and stressed the fact that even when it did, its effects tended to be short lived (Skinner, 1953). Of course, we recognize that our results are only directly applicable to procedural learning paradigms and may well not generalize to the very wide range of behaviors for which reward and punishment are used as modulators. Nevertheless, procedural learning is an important aspect of learning, and the data we present might well be applicable to the rehabilitation of motor function in patients with various forms of motor disability including those after a stroke.

The second question is whether our results tell us anything about the interaction among positive and negative incentives (reward and punishment), motivation, and learning. Although this issue has been dealt with extensively in the literature on Pavlovian conditioning (Dickinson and Balleine, 2002) it has not been clear whether reward and punishment interact with separate motivational systems that have different neurochemical and neuroanatomical substrates and produce differential behavioral effects (Bindra, 1974; Dickinson and Dearing, 1979; Dayan and Balleine, 2002), and the question has never been addressed in the context of implicit procedural learning. The behavioral task we used enabled us to separate the effects of reward and punishment on motor performance from those on motor learning. The fact that there were qualitatively distinct effects on these behavioral measures suggests that reward and punishment may actually engage qualitatively different motivational systems. However, it is not clear whether there were quantitative differences in motivation between these groups. The similarity of the RT-error trade-off in both groups (Fig. 1) might be interpreted as indicating similar motivation; however, a more parsimonious explanation is that they used similar criteria for this trade-off. The negativity bias (Taylor, 1991) in the decision-making literature would suggest that the motivation might be stronger in the punished group, given the equal monetary value of reward and punishment, making it all the more remarkable that the rewarded subjects clearly learned more. If reward and punishment actually access separate motivation systems, then one would expect that this should be evident in the underlying neural substrate. In the current study, the activity in the dorsal and ventral striatum that related to the reward per se replicates previous findings (McClure et al., 2003; O’Doherty et al., 2003, 2004) and most likely represents the neural correlate of the dopaminergic neurons coding a prediction error signal in addition to being consistent with the two process account of reinforcement learning (Montague et al., 1996; Sutton and Barto,

1998). In contrast, punishment led to activation predominantly in the inferior frontal gyrus and in the insula, the latter being the most consistently activated area in a variety of studies relating to punishment (Elliott et al., 2000; Sanfey et al., 2003; Daw et al., 2006). It has been hypothesized that punishment, even when it is associated with striatal activation (Seymour et al., 2004), does not operate through the dopaminergic system (Ungless et al., 2004) but is more likely mediated through the serotonergic system originating in the median raphe nucleus (Daw et al., 2002). The end result of activating such a motivational system in our case was the change in performance we documented with which the insula was the only area to be significantly correlated. The difference between the results of this and many other studies is that we were able to correlate the neural substrates of reward and punishment with qualitatively different behavior outcomes suggesting that these modulators might indeed operate through different motivational systems.

The final issue to consider is how the results contribute to our knowledge of the role of the basal ganglia in procedural learning. There is a general consensus that the basal ganglia are an important substrate for procedural learning (Grafton et al., 1995; Rauch et al., 1997; Willingham et al., 2002), particularly when learning becomes more established (Poldrack et al., 2005; Seidler et al., 2005). The location of activation associated with learning in the current study, the dorsal striatum (putamen), is similar to that identified by others (Poldrack et al., 2005; Seidler et al., 2005). The dorsal striatum is involved in learning stimulus–action–reward associations during instrumental learning (Haruno et al., 2004; O’Doherty et al., 2004). In our experiment, the association between stimulus and action was deterministic; therefore, the activity in the putamen cannot be related to learning the stimulus–action association. Similarly, because activity in the putamen was higher in the sequence compared with the random blocks regardless of reward delivery, this activity is not solely related to action–reward association. Our data suggest that reward facilitates procedural motor learning within the motor system by modulating the activity of the putamen which has extensive connections to premotor areas. It is likely that the effect is translated into an improvement in motor learning by a dopamine-induced potentiation of corticostriatal synapses in the striatum, similar to that which occurs after direct stimulation of the substantia nigra (Reynolds et al., 2001). In the same context, our results may also be relevant to patients with Parkinson disease, who show some deficits in procedural learning tasks, but who are disabled primarily because of an inability to produce coherent sequences of over-learned movements. It is possible that the lack of dopamine in these patients results in an impairment of an intrinsic reward system (Mazzoni et al., 2007), based on an internal representation of motor performance, thus disrupting the type of corticostriatal facilitation we demonstrated, and thereby affecting the performance of sequential movements.

References

- Berns GS, Sejnowski TJ (1998) A computational model of how the basal ganglia produce sequences. *J Cogn Neurosci* 10:108–121.
- Bindra D (1974) A motivational view of learning, performance, and behavior modification. *Psychol Rev* 81:199–213.
- Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P (2001) Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30:619–639.
- Daw ND, Kakade S, Dayan P (2002) Opponent interactions between serotonin and dopamine. *Neural Netw* 15:603–616.
- Daw ND, O’Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879.
- Dayan P, Balleine BW (2002) Reward, motivation, and reinforcement learning. *Neuron* 36:285–298.
- Delgado MR (2007) Reward-related responses in the human striatum. *Ann N Y Acad Sci* 1104:70–88.
- Delgado MR, Locke HM, Stenger VA, Fiez JA (2003) Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn Affect Behav Neurosci* 3:27–38.
- Dickinson A, Balleine BW (2002) The role of learning in the operation of motivational systems. In: *Stevens’ handbook of experimental psychology: learning, motivation, and emotion*, Ed 3 (Gallistel R, ed), pp 497–533. New York: Wiley.
- Dickinson A, Dearing MF (1979) Appetitive-aversive interactions. In: *Mechanisms of learning and motivation* (Dickinson A, Boakes RA, eds), pp 203–231. Hillsdale, NJ: Erlbaum.
- Elliott R, Friston KJ, Dolan RJ (2000) Dissociable neural responses in human reward systems. *J Neurosci* 20:6159–6165.
- Frensch PA, Wenke D, Rüniger D (1999) A secondary tone-counting task suppresses expression of knowledge in the serial reaction time task. *J Exp Psychol Learn Mem Cogn* 25:260–274.
- Grafton ST, Hazeltine E, Ivry R (1995) Functional mapping of sequence learning in normal humans. *J Cogn Neurosci* 7:497–510.
- Haruno M, Kuroda T, Doya K, Toyama K, Kimura M, Samejima K, Imamizu H, Kawato M (2004) A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J Neurosci* 24:1660–1665.
- Houk JC, Adams JL, Barto AG (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: *Models of information processing in the basal ganglia* (Houk JC, Davis JL, eds), pp 249–270. Cambridge, MA: MIT.
- Mazzoni P, Hristova A, Krakauer JW (2007) Why don’t we move faster? Parkinson’s disease, movement vigor, and implicit motivation. *J Neurosci* 27:7105–7116.
- McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346.
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947.
- Nissen MJ, Bullemer P (1987) Attentional requirements of learning: evidence from performance measures. *Cognit Psychol* 19:1–32.
- O’Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–454.
- O’Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337.
- Packard MG, Knowlton BJ (2002) Learning and memory functions of the basal ganglia. *Annu Rev Neurosci* 25:563–593.
- Pavlov IP (1927) *Conditioned reflexes*. London: Routledge and Keegan Paul.
- Poldrack RA, Sabb FW, Foerdes K, Tom SM, Asarnow RF, Bookheimer SY, Knowlton BJ (2005) The neural correlates of motor skill automaticity. *J Neurosci* 25:5356–5364.
- Rauch SL, Whalen PJ, Savage CR, Curran T, Kendrick A, Brown HD, Bush G, Breiter HC, Rosen BR (1997) Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. *Hum Brain Mapp* 5:124–132.
- Reynolds JN, Hyland BI, Wickens JR (2001) A cellular mechanism of reward-related learning. *Nature* 413:67–70.
- Robertson EM (2007) The serial reaction time task: implicit motor skill learning? *J Neurosci* 27:10073–10075.
- Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. *Science* 300:1755–1758.
- Schultz W (2002) Getting formal with dopamine and reward. *Neuron* 36:241–263.
- Seidler RD, Purushotham A, Kim SG, Ugurbil K, Willingham D, Ashe J (2005) Neural correlates of encoding and expression in implicit sequence learning. *Exp Brain Res* 165:114–124.
- Seymour B, O’Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ,

- Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667.
- Skinner BF (1938) *The behavior of organisms*. New York: Appleton-Century.
- Skinner BF (1953) *Science and human behavior*. New York: Macmillan.
- Squire LR (2004) Memory systems of the brain: a brief history and current perspective. *Neurobiol Learn Mem* 82:171–177.
- Suri RE, Schultz W (1998) Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Exp Brain Res* 121:350–354.
- Sutton RS, Barto AG (1998) *Reinforcement learning*. Cambridge, MA: MIT.
- Taylor SE (1991) Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychol Bull* 110:67–85.
- Thorndike EL (1911) *Animal intelligence*. New York: Macmillan.
- Ungless MA, Magill PJ, Bolam JP (2004) Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 303:2040–2042.
- Willingham DB, Goedert-Eschmann K (1999) The relation between implicit and explicit learning: evidence for parallel development. *Psychol Sci* 10:531–534.
- Willingham DB, Wells LA, Farrell JM (2000) Implicit motor sequence learning is represented in response locations. *Mem Cogn* 28:366–375.
- Willingham DB, Salidis J, Gabrieli JD (2002) Direct comparison of neural systems mediating conscious and unconscious skill learning. *J Neurophysiol* 88:1451–1460.